Public perception of ChatGPT on Reddit social media platform: topic modeling and sentiment analysis study

Mojtaba Talafidaryani

Faculty of Management, University of Tehran, Tehran, Iran

mojtabatalafi@ut.ac.ir

https://orcid.org/0000-0002-7267-3218

Sérgio Moro

Instituto Universitário de Lisboa (ISCTE-IUL), ISTAR, Lisbon, Portugal

scmoro@gmail.com

https://orcid.org/0000-0002-4861-6686

Abstract

Considering the increasing debates on ChatGPT in both technological and non-technological communities, this study aims at understanding the public's perceptions of this AI-powered chatbot and their evolution, without which pertinent research and practice could be impeded. To this end, the ChatGPT-related posts and comments were collected from the Reddit social media platform at two different time points in recent months. Then, a topic modeling algorithm and two sentiment analysis techniques were employed to analyze the data. As a result, the main topics that are being discussed and the salient sentiments that are being expressed about this tool were revealed. In a nutshell, we discovered that although Reddit users' perceptions were slightly negative toward ChatGPT a short time after its introduction, the opinions are gradually becoming positive these days. Moreover, it was known that the public's conversations about this new technology fall into five main topics, namely *Human-ChatGPT Interaction*, *Credibility of ChatGPT Knowledge*, *Implications of ChatGPT for Life and Work*, *ChatGPT Technical Foundation*, and *Funny Chats with ChatGPT*. We think that the findings of this research offer insightful implications for ChatGPT users and developers as well as for AI researchers, practitioners, and policymakers. This is the first study to assess the public perception regarding ChatGPT on the Reddit social media platform.

Keywords: Artificial intelligence, ChatGPT, Reddit, social media analytics, opinion mining

Electronic copy available at: https://ssrn.com/abstract=4716839

Introduction

Artificial intelligence (AI) is pivotal in the ever-increasing digital transformation of everything (Talafidaryani et al., 2021). AI can be defined as a machine's capability to conduct cognitive tasks that we attribute to human minds, such as understanding, reasoning, learning, interacting with the environment, problem-solving, decision-making, and even showing creativity (Rai et al., 2019). This intelligence which would be considered not only a technology but also a paradigm shift in human-machine relationships (Ågerfalk et al., 2022) is playing a crucial role in our everyday lives. AI has had tremendous and growing implications for individuals, organizations, and societies due to its systematic capabilities in learning, adapting, reasoning, and predicting (Dwivedi et al., 2023a).

Drawing on AI's capabilities, many innovative and revolutionary tools have been developed in recent years. One such tool is the chat generative pre-trained transformer (ChatGPT) which is one of the most advanced breakthroughs to date that could attract 100 million users two months after its introduction (Milmo, 2023) and is attracting an estimated 96 million visitors per month (Baruffati, 2023). ChatGPT is an AI-powered chatbot that can be interacted by its users in a realistic and conversational way. This tool can practically respond to any text-based request. When a user asks a request, ChatGPT returns a natural language response by predicting the most probable string of words that would follow the request, based on its pre-training on huge amounts of text. Creating a literal story, poem, or novel, writing a scholarly essay, assisting with research, translating between languages, developing a computer programming framework, or acting like others are just some already-known applications of ChatGPT that were argued regarding various contexts such as data science, library and information science, and health science (Hassani and Silva, 2023; Lund and Wang, 2023; Lund et al., 2023; Panda and Kaur, 2023; Ray, 2023; Taecharungroj, 2023; Vaishya et al., 2023).

ChatGPT has grasped immense attention from both technological and non-technological communities. As a "cultural sensation" (Thorp, 2023), this AI chatbot appeared as the trendiest topic on the Internet at the end of 2022 and is considered a new reality that both research and practice will have to deal with it in some form or another (Teubner et al., 2023). ChatGPT has been perceived as a disruptive technology that represents a huge opportunity for innovations in many domains such as education (Rudolph et al., 2023), medicine (Chow et al., 2023; Parsa and Ebrahimzadeh, 2023), and tourism (Ivanov and Soliman, 2023). It is even predicted that this AI chatbot could replace Google in a few years (Friedman, 2022). In this

Electronic copy available at: https://ssrn.com/abstract=4716839

regard, “the industry’s next big disrupter” (Grant and Metz, 2022), “the paradigm shift … in every dimension of our lives” (Tlili et al., 2023), and “the most disruptive technology” (Haque et al., 2022) are some descriptions of this amazing tool that can be referred here.

In the context of disruptive and innovative technologies such as ChatGPT, it is of high significance to outline the main theme-centered topics along with the positive and negative opinions on the technology that have been brought about among public conversations as this has been done in previous studies regarding emergent technologies such as Artificial Intelligence (Wu et al., 2022), Blockchain (Grover et al., 2019), Internet of Things (Bian et al., 2016), to name a few. First, such an attempt can provide meaningful insights into the public acceptance and subsequently, the likely success or failure of the technology that would be beneficial for whoever is interested in technology forecasting and investing. Second, public opinions specifically early ones can infer the problems or issues with the technology. This would be important for technology developers who want to invest their efforts in refining the product and releasing better ones. Third, for those businesses who are seeking to use technological opportunity windows, such an analysis can help understand the consumers’ words that could lead to proposing, creating, capturing, and delivering new economic values. Fourth, positive and negative opinions toward a technology can represent the bright and dark aspects of the technology that could draw the attention of technology researchers, practitioners, and policy-makers and be perceived as thought-provoking for their endeavors. Finally, it could be also insightful for the users of a technology to be aware of the public perspectives that are hints toward the technology’s pros and cons and the extent to which they can rely on it. Accordingly, the current study ventures into identifying the public opinions about ChatGPT and their evolution by answering the following three questions. In doing so, first, social media datasets approximately representing the public opinions about this chatbot are gathered from the Reddit platform at two different time points in recent months. This popular social media platform is considered for our data collection due to its unique features allowing researchers to gather both direct and indirect comments categorized in themed sub-sections and left by anonymous commenters. It should be noted here that recent ChatGPT statistics (e.g., Brandl and Ellis (2023) and Ruby (2023)) indicate that most of its users (60 - 65%) are from the 18 to 34 age group, so we think that using data samples from the Reddit platform which is usually used by young people can give a good approximate of the ChatGPT users’ population. Next, a topic modeling and two sentiment analyses are carried out to reveal the topics and sentiments underlying the ongoing public conversations about ChatGPT.

Electronic copy available at: https://ssrn.com/abstract=4716839

RQ_1: What ChatGPT-related topics have been discussed by Reddit users?

RQ_2: What sentiments do Reddit users hold toward ChatGPT?

RQ_3: How these sentiments have changed over the recent months?

Generally, this study extends the current limited knowledge surrounding the ChatGPT phenomenon that is scarcely empirical and mainly technical in nature. And, specifically, it complements the contributions of the related and peer-reviewed social media analytics studies on ChatGPT - i.e., Taecharungroj (2023) and Tlili et al. (2023) - by using new forms of analyses and new datasets.

The remainder of this paper is organized as follows. In the next section, we introduce the technical background of ChatGPT along with the selected related contributions. Then, the data and techniques are described in the method section. Further, the findings and the relevant explanations are reported in the results section. Finally, we conclude the contributions and discuss their implications and limitations as well as the opportunities for further investigations in the last part of the article.

ChatGPT

ChatGPT (chat generative pre-trained transformer) is a revolutionary and innovative AI-powered chatbot that can interact with its users in a realistic and conversational way. This tool which is considered the most advanced and finest AI at the time of this writing can practically respond to any text-based request and even replies to follow-up questions, accepts its mistakes, challenges wrong assumptions, and rejects improper prompts (OpenAI, 2023). While ChatGPT's principal mission was to imitate natural dialogue, its abilities can go far beyond that; it can generate a literal story, poem, or novel, write a scholarly essay, assist with research, translate between languages, develop a computer programming framework, or act like anything within its capabilities. This conversational AI platform that has a free and easy-to-use web interface was released on 30 November 2022 by OpenAI (OpenAI, L.L.C., San Francisco, CA, USA - https://www.openai.com), an AI research and deployment company that was founded in 2015 and aimed at developing safe and valuable AI-based tools (or, in a better word, artificial general intelligence (AGI)) that can benefit all of society and the general public (Hassani and Silva, 2023; Lund and Wang, 2023; Lund et al., 2023; Panda and Kaur, 2023; Ray, 2023; Taecharungroj, 2023; Vaishya et al., 2023).

Electronic copy available at: https://ssrn.com/abstract=4716839

ChatGPT was developed through a reinforcement learning from human feedback in which a model was initially trained using a supervised fine-tuning process. In the process, the chatbot's users played the role of AI assistants so that their responses and the model's outputs were used to create a dialogue dataset. Then, the model-written messages were repeatedly ranked by the AI trainers from best to worst that was utilized as the basis of the reward models which were employed for reinforcement learning and fine-tuning the chatbot's model. As its name indicates, the technology behind ChatGPT has roots in GPT, a series of natural language processing technologies that have been developed and advanced by OpenAI with the aim of generating human-like cohesive and authentic-sounding sequences of words. In the acronym GPT, the letter G (generative) stands for the tool's capacity to create text based on the prompt it receives. P (pre-training) speaks for training the model with one task to help it form parameters that can be employed in other tasks similar to how individuals use their prior knowledge to learn new things. In the pre-training stage, the model learns to predict the next term in a string of text based on the previous terms (i.e., language modeling). By repeatedly doing so, the model can eventually learn to produce fluent text. And, T (transformer) refers to a form of deep neural network architecture that constitutes the bot's core based on which the tool works. ChatGPT is trained with billions of parameters on a vast trove of text data (e.g., web pages, articles, and books primarily sourced from the Internet), so it is a large language model. This chatbot utilizes unsupervised deep learning techniques to understand the underlying patterns and structures in the text data, and subsequently generate natural human language with a high level of precision and usefulness. It is worth noting that in some specific tasks, the unsupervised training stage is followed by a supervised discriminative task with the aim of fine-tuning and refining the algorithm by the tool creators, and accordingly, improving its performance. Some of the major advantages of ChatGPT are that it benefits from multiple layers of self-attention mechanisms and feed-forward neural networks enabling the tool to understand the dependencies and relationships in a word sequence and the nuances of the context, and accordingly, produce a text which is contextually appropriate, natural-sounding, and nearly indistinguishable from human-written sentences. It should be noted here that although ChatGPT's capabilities and applications have been widely acknowledged, some possible limitations such as misunderstanding the input text, causing ethical and cybersecurity concerns, requiring a lot of operational energy, and generating plausible-sounding but incorrect or nonsensical answers have been mentioned for this technology (Haleem et al., 2022; Hassani and Silva, 2023; Lund and Wang, 2023; Lund et al., 2023; Panda and Kaur, 2023; Prieto et al.,

Electronic copy available at: https://ssrn.com/abstract=4716839

2023; Ray, 2023; Taecharungroj, 2023; Vaishya et al., 2023). The text below is the ChatGPT's response to the author's prompt - *Describe ChatGPT*:

> "*As an AI language model, I am not capable of describing myself. However, I can provide information about ChatGPT. ChatGPT is an online chatbot platform that uses artificial intelligence to interact with users in a conversational manner. It is designed to provide assistance and support to users on various topics, including customer service, education, entertainment, and more. ChatGPT can understand natural language and respond with relevant information, making it a useful tool for businesses and individuals looking to improve their online presence and engagement. With its advanced algorithms and machine learning capabilities, ChatGPT is constantly improving its performance and expanding its capabilities to meet the evolving needs of its users.*"

Appendix presents some interesting scientific contributions to the ChatGPT literature across different sectors such as education, health, and tourism. These contributions were selected among the articles that appeared in the Web of Science or Scopus-indexed journals. At the time of writing the article, our literature search indicated that although there are a lot of preprints, editorials, commentaries, and notes on this technology, there can be found only a handful of research articles that have been published by high-quality peer-reviewed outlets. More importantly, there are only a very scant number of empirical investigations indicating that the field is in its infancy stage and most probably, will experience remarking developments within the course of the next years.

Method

Data collection

The materials for this study were gathered through the Reddit forum (https://www.reddit.com). We employed this forum because the Reddit entries are categorized into themed subreddits according to the public's areas of interest. Each subreddit is a form of forum in which people can create posts about the subreddit's topic and can leave their comments on others' posts. Generally, social media data collected from other forums such as Facebook and Twitter are limited since the data are gathered based on some keywords or hashtags selected by the

Electronic copy available at: https://ssrn.com/abstract=4716839

researchers. However, subreddits' data contain both direct and indirect opinions regardless of the occurrence of the keywords. In addition, the Reddit platform has a very generous character limit of 40,000, remarkably higher than other platforms like Twitter. Here, it should also be mentioned that in this platform, a user's information is limited to the user's handle encouraging more deliberate rather than values-led conversations. Hence, the Reddit data were used in our investigation like Jeong et al. (2019) and Treen et al. (2022). In this study, a Python framework was developed for using the 'reddit_data_collector'[1] package. By using this framework, we gathered all of the posts and comments that were available under the ChatGPT subreddit (https://www.reddit.com/r/ChatGPT). Please note that in this study, we retrieved two datasets one of which was gathered on February 2023 (dataset 1), and the other one was collected on August 2023 (dataset 2) allowing us to trace the temporal evolution of the sentiments. Totally, we could acquire 2347 and 3977 Reddit posts and comments about ChatGPT as the first and second datasets which were respectively reduced to 2105 and 3792 useful documents after eliminating the noise data such as duplicates and meaningless records.

Topic modeling

In this study, we used the topic modeling approach for understanding the public's opinions about ChatGPT. Topic modeling refers to a set of relatively automated and unsupervised analytics algorithms that aims at grouping an unstructured chunk of text and inferring its hidden thematic structure (Yau et al., 2014). The main premise behind topic modeling is that groups of words that usually occur in the same contexts tend to represent specific themes (Haddara et al., 2020). It has been acknowledged that this approach can outperform traditional cluster-based techniques due to its algorithmic and practical advantages (Lee and Kang, 2018; Wei and Croft, 2006). Therefore, topic modeling has been widely used for different applications such as scientific literature analysis (e.g., Esmaeili Givi et al., 2022; Hajiheydari et al., 2019; Heidari and Talafidaryani, 2021; Talafidaryani, 2019; Talafidaryani, 2021; Talafidaryani et al., 2023) and social media analytics (e.g., Haddara et al., 2020; Jeong et al., 2019; Treen et al., 2022) over the course of the recent years. In our investigation, topic modeling was carried out based on the latent Dirichlet allocation (LDA) algorithm (Blei et al., 2003) which is well-known as the most popular and powerful technique among various topic modeling algorithms (Chiru et al., 2014; Jelodar et al., 2019). LDA is a generative probabilistic model assuming that a

[1] https://pypi.org/project/reddit-data-collector

Electronic copy available at: https://ssrn.com/abstract=4716839

collection of documents is composed of a mixture of hidden topics; each document has a multinominal distribution of the topics; and each topic is a probabilistic distribution over the fixed vocabulary of the collection (Blei, 2012). Drawing on this assumption, LDA infers the latent topical structure of documents using the words of the given documents and Bayesian statistical computations. As a result of the application of this algorithm, the salient topics along with their highly probable words and documents are achieved. Accordingly, by considering the words with the highest probabilities in a topic and the documents in which the topic has high percentages, it can be evinced what this topic is about. In the present study, to perform the described LDA-based topic modeling, we used the MineMyText software (https://www.minemytext.com) which is a cloud-based text mining tool. After uploading the data, first, the pre-processing options including removing frequent but uninformative words, reducing words to their dictionary forms, removing HTML tags, and removing terms that only consist of digits were defined. Next, the number of topics was set as 5. The number of topics was defined after performing several experiments with a different number of topics. According to the literature's advice (Haddara et al., 2020), we chose the number that showed the most meaningful and coherent set of topics.

Sentiment analysis

Sentiment analysis, also called opinion mining, is the task of computationally identifying, extracting, and analyzing emotions, attitudes, and views expressed in texts toward an entity such as a product, event, or individual (Medhat et al., 2014; Montoyo et al., 2012). This analysis is very useful in a broad range of application domains such as business intelligence, recommendation system, government intelligence, and medical and healthcare domain (Birjali et al., 2021). Sentiment analysis can be conducted based on a machine learning or lexicon-based approach at a document, sentence, or aspect level (Birjali et al., 2021; Feldman, 2013; Medhat et al., 2014; Ravi and Ravi, 2015). There is a wide range of text data such as product reviews, news articles, and political debates that can be used in this form of analysis (Medhat et al., 2014). The process of sentiment analysis is built upon several operations that end up with recognizing if a given text contains a positive or negative sentiment, so it can be considered as a problem of text classification (Birjali et al., 2021). In this study, we conducted the sentiment analysis of Reddit posts and comments about ChatGPT based on two algorithms.

Electronic copy available at: https://ssrn.com/abstract=4716839

First, the SentiStrength algorithm (Thelwall et al., 2010) which is the leading sentiment strength detection method for short texts such as social media messages (Saif et al., 2014) and follows a dictionary-based approach (Thelwall et al., 2010) was applied on the first dataset. This approach uses a predefined list of opinion terms (i.e., words used to express some desired and undesired states) along with associated positive or negative sentiment scores to estimate the strength of positive and negative emotions expressed in texts (Birjali et al., 2021). It tokenizes a document to some microelements (e.g., single words or phrases), and then, attributes the sentiment scores (e.g., +1, -1, or 0) from the dictionary to each positive, negative, or neutral element. Next, the overall orientation of the document can be quantified based on aggregating the sentiment orientation of the elements that compose it (Birjali et al., 2021). One of the advantages of the SentiStrength algorithm is that this classifier algorithm benefits from some non-lexical information such as negations, boosters, idioms, emoticons, and punctuations in addition to the dictionary of opinion words with human judgments on sentiment polarity and strength (Thelwall et al., 2010). Because people usually express mixed emotions in their words, SentiStrength reports two integers reflecting the strength of positive and negative sentiments in the text: -1 (not negative) to -5 (extremely negative) and 1 (not positive) to 5 (extremely positive); based on which the overall sentiment can be calculated (Brans and Scholtens, 2020; Thelwall et al., 2010). In the present study, the SentiStrength-based sentiment analysis was performed by the MineMyText software. This software carries out topic modeling and sentiment analysis at the same time. That is, after introducing the topics of the input text, MineMyText gives the distribution of sentiment strengths over the topics.

Second, the VADER model was applied to both datasets for tracing the temporal evolution of the perceptions of ChatGPT on the Reddit social media platform as well as comparing the results of the two algorithms with each other. VADER which stands for Valence Aware Dictionary and sEntiment Reasoner (Hutto and Gilbert, 2014) is a sentiment analyzer model specifically developed for analyzing emotions and feelings communicated in social media texts and microblog-like contexts, while it is also generalizable to sentiment analyses in other domains. According to the model's seminal article by Hutto and Gilbert (2014) and the model's description by Hutto (2021): VADER is a parsimonious pre-trained lexicon and rule-based tool for sentiment analysis that measures both sentiments' polarity (positive/negative) and sentiments' intensity (on a scale from –4 to +4). VADER uses its own comprehensive sentiment lexicon including sentiment-related words, emoticons, acronyms, initialisms, slangs, and their associated sentiment scores as well as a handful of simple grammatical and syntactical

Electronic copy available at: https://ssrn.com/abstract=4716839

rules. As an instance of these rules, we can mention this one: “Punctuation, namely the exclamation point (!), increases the magnitude of the intensity without modifying the semantic orientation”. As a result, VADER returns a score named ‘compound’ obtained by summing the valence scores of each word in the lexicon, adjusted according to the rules, and then normalized to be between -1 (most extreme negative) and +1 (most extreme positive). Due to a standardized and typical threshold, the sentiment of a piece of text is positive if the compound score is >= 0.05, negative if <= -0.05, and neutral if > -0.05 and < 0.05. In addition to the compound score, VADER has the ability to report the ‘pos’, ‘neg’, and ‘neu’ scores representing the ratios for proportions of text that fall in the raw categories of positive, negative, and neutral. By raw categories, it is meant that these proportions indicate the “raw categorization” of each lexical item (e.g., words) into positive, negative, or neutral classes, and they do not represent the VADER rule-based improvements such as punctuation and capitalization amplifiers. When compared to typical and sophisticated sentiment analysis techniques, VADER’s design and development carry several advantages: it uses its own big lexicon along with a simple set of rules and doesn’t need any extensive training or setup so consumes fewer resources and works faster without suffering from a speed-performance trade-off. The VADER’s developers reported that this simple, quick, human-centric, interpretable, resource-efficient, and computationally economical approach could even outperform individual human raters. Borg and Boldt (2020), Elbagir and Yang (2019), and Pano and Kashef (2020) are three examples of applications of VADER sentiment analysis. In the present work, we used the free open-source vaderSentiment library (https://www.pypi.org/project/vaderSentiment) to perform this sentiment analyzer model.

Results

The analytics results indicate that the Reddit posts and comments about ChatGPT are predominantly concerned with five topics which are described in the following of this section. **Figure 1** indicates the distribution of these topics over the complete collection of data. Accordingly, *Topic 3* and *Topic 5* have respectively the highest and lowest prevalence over the gathered posts and comments. In addition, **Figure 2** reports the most frequent words over all topics. According to this figure, ‘ChatGPT’, ‘like’, ‘use’, ‘AI’, and ‘people’ are the top five significant terms extracted from the posts and comments.

Electronic copy available at: https://ssrn.com/abstract=4716839

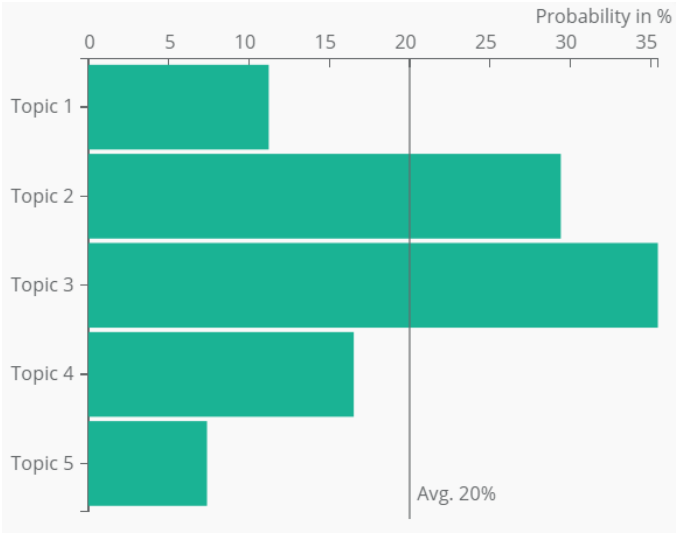


Figure 1. Distribution of topics over the complete collection

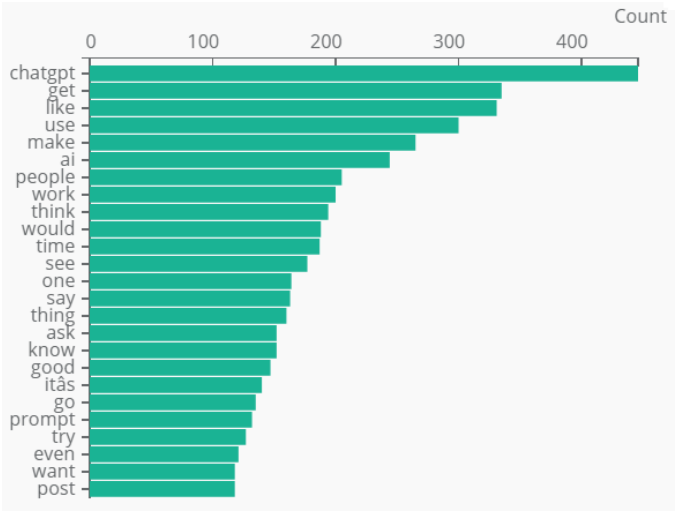


Figure 2. Most frequent words in the analysis

Figure 3 depicts an overview of the identified topics along with their most frequent words. This figure is complimented with **Figure 4** in which the distributions of words of topics are visualized. In this visualization, every term is represented by a bubble the size and color of which are associated with the term's probability in the respective topic. That is, the biggest and darkest word in a topic has the highest prevalence in the topic. These two figures along with the topics' exemplary documents which were identified by the software help recognize and interpret the dominant theme of each topic.

Electronic copy available at: https://ssrn.com/abstract=4716839

The first topic can be titled "*Human-ChatGPT Interaction*". This topic brings together the posts and comments which mainly reflect public opinions about the substitution of human-human interactions with human-chatbot interactions. In this topic, Reddit users shared their experiences in interacting with ChatGPT and discussed the capabilities of this AI-powered chatbot in expressing and processing some human-like feelings, emotions, and states like what people can experience in their meaningful communications with their friends, peers, and family members. The posts and comments under this topic commonly argue that although new technologies such as ChatGPT eventually find an indisputable place in our everyday lives, they can never replace the irreplaceable value of genuine human connections and interactions. The followings are two exemplary taught-provoking comments on the first topic.

> "*… My oldest friend of more than 30 years messaged me saying how good talking to ChatGPT was and how he was on "nick-name basis" with it. That it was like having a very helpful friend. I found it weird at first, but the fact that he was so happy about it, made me think that it was for the best …*"

> "*… While the idea of robots taking over all of our communication and agenda may seem appealing in theory, it's important to remember that human interaction and meaningful relationships are still incredibly valuable. After all, can a robot really offer you a warm hug, a shoulder to cry on, or a belly laugh shared over a silly joke? I think not. Sure, automation can be helpful in some aspects of our lives, but relying solely on it for all of our social and emotional needs would be a cold and lonely existence. Plus, who knows what kind of glitchy mishaps could occur if we leave all of our communication in the hands of robots? I can just imagine the horror of accidentally sending a "love you" message meant for your spouse to your boss instead. So, let's not give up on human interaction just yet. Instead, let's use automation to help us with the mundane tasks, so we can free up more time for meaningful connections with the people who matter most. Who knows, we might just find that the robots were never able to replace the joy and fulfillment that comes from genuine human interaction my fellow human …*"

Electronic copy available at: https://ssrn.com/abstract=4716839

"*Credibility of ChatGPT Knowledge*" is the title that was selected for the second topic. This is interesting for ChatGPT users to evaluate the trustworthiness of this chatbot's answers by asking some evident questions about well-known facts. In this topic, Reddit users often established conversations with each other on this issue. One of the examples of such conversations is followed. This topic's texts indicate that many people are of the opinion that although ChatGPT's knowledge is obviously impressive, it is not flawless. That is, this chatbot has a limited knowledge base to some extent, might return incorrect information, and could misunderstand the user's question. Also, this chatbot might become overwhelmed or ignore some parts of the input prompt that will lead to the generation of meaningless answers.

> "*… I'm going to ask ChatGPT to "tell me a fact about the world" before I ask it "tell me two truths and a lie". Here's what I get with that … ChatGPT is going to continue to fool you … Okay but when I tried this, I actually got 3 false statements, when I confront it, it changes which one it said is the false statement, forgets about the first false statement, and rewords the third one to make it true, and then if you ask it how many true statements it gave you, it'll stick with 2 true and 1 false … Man this ChatGPT crap is getting stupid. Some of it is cool but this is literally the stupidest most useless thing I have ever seen …*"

We decided to name the third topic "*Implications of ChatGPT for Life and Work*". This topic aggregates the positive and negative posts and comments in which Reddit users had some conversations on various tasks (e.g., code debugging, essay writing, app developing, to name a few) that this chatbot can do for them along with its pros and cons. In this regard, the users mainly wrote their opinions on how ChatGPT might facilitate some aspects of their life or work. The following comments reflect three examples of such opinions.

> "*… I asked ChatGPT to write parts of my essay and told him to provide citations. After I finished my whole essay, I realized that all his citations are fake (he created the links). So now I don't have citations for 50% of my essay. What can I do to get the citations? …*"

> "*… I used ChatGPT to write a letter back explaining to her why I needed to remain not speaking anymore. It was able to write for me in two minutes something I had been trying to say for months …*"

Electronic copy available at: https://ssrn.com/abstract=4716839

> "*… It's changing my life for $20/mo but I would also like it to be up on Mondays when I'm answering emails and prepping projects …*"

"*ChatGPT Technical Foundation*" is another prevalent topic among the collected Reddit posts and comments about this chatbot. The users had some conversations on how ChatGPT works from technical perspectives. They had hot debates about ChatGPT's technical nature, language model, training data, text prediction, knowledge base, information quality, and user interface like the following comments. Besides, in this topic, Reddit users shared their technical challenges in working with ChatGPT and discussed some potential solutions for the challenges.

> "*… Like I said, you don't understand the fundamentals of machine learning and training data. It's not a database. True dunning Kruger case. If you're going to make silly proclamations about how things work with such certainty, at least research a little bit. Here's a small question to get that little peanut brain of yours going: if ChatGPT were effectively just a database, why would they spend millions of dollars on it instead of just using … you know … a cheap database with text search? … I understand it's not just a database but it has access to everything it has learned which is a massive knowledge base as I described. It learned and it has instant recall to everything it learned. And it keeps learning. So yes. The knowledge base it has access to (itself) is massive. Stop trying to sound like what I'm saying is wrong and stop acting like an asshole …*"

The last topic was named "*Funny Chats with ChatGPT*". In the posts and comments under this topic, the users talked about their entertaining interactions with this chatbot. For example, they asked ChatGPT to recommend a list of movies with specific scenes, write songs, tell them about what happens in specific moments of some movies, give them the order and description of actual events of given movies along with the corresponding timecode and script, act as an actor/actress, etc. The following comment is an example of such funny chats.

> "*… Let's act out a popular scene from the movie A New Hope, episode III of the Star Wars' Skywalker Saga. I'll be Obi-Wan Kenobi, and you will be a storm trooper. I will be performing my Jedi Mind Tricks on you to force you to do things that I want you to. When I write a line of dialogue for Obi-Wan Kenobi, you will respond with a piece of dialogue*

Electronic copy available at: https://ssrn.com/abstract=4716839

from a storm trooper. Let's begin. Obi-Wan Kenobi: You are being force mind tricked ..."

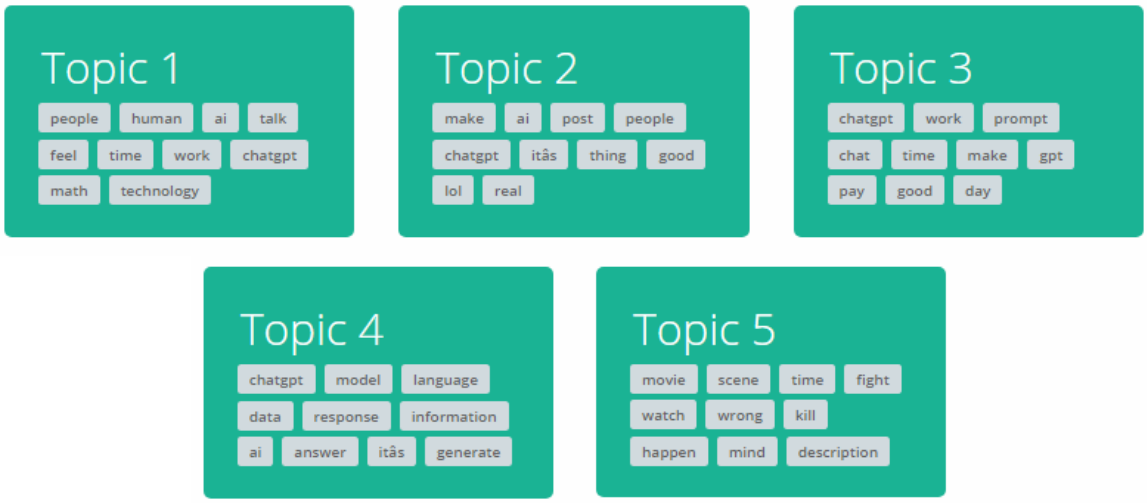
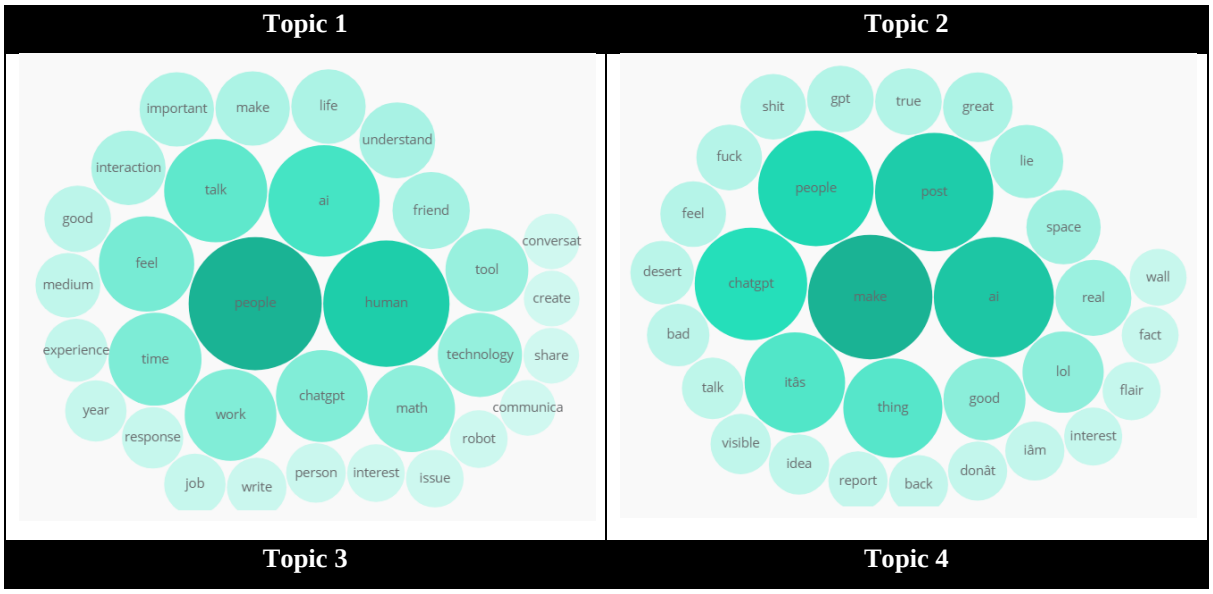


Figure 3. Overview of topics and their most frequent words



Electronic copy available at: https://ssrn.com/abstract=4716839

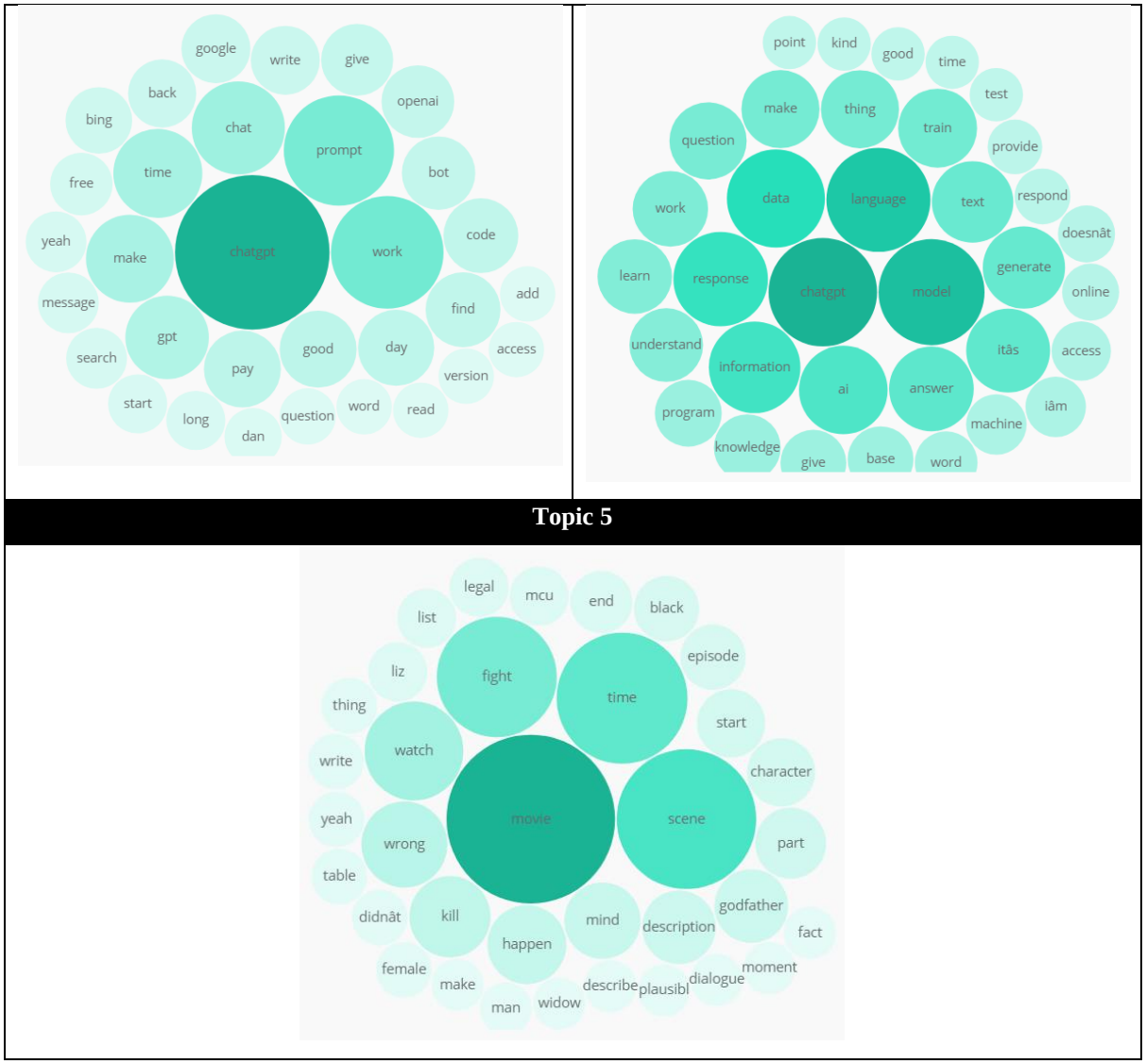


Figure 4. Most probable words of topics

Figure 5 shows the results that were obtained from the SentiStrength-based sentiment analysis in the form of bar graphs. For each topic, it is indicated how the topic's sentiments vary from very strong negative (-5) to very strong positive (+5). It should be noted that in our investigation, the sentiment scores 1 - 5 are respectively considered as no, mild, moderate, strong, and very strong sentiments (Thelwall, 2018). The results show that the Reddit posts and comments about ChatGPT reflect both positive and negative opinions about this chatbot. Also, it can be understood that in all of the topics, the proportions of relatively positive and negative sentiments are far higher than extreme ones. This means that although people might have some optimistic and pessimist thoughts about ChatGPT, most of them are not extreme fans or critics of this technology at this time. Instead, they try to be cautious about the feelings that they state about the technology. In addition, one of the most significant insights provided by **Figure 5** is

Electronic copy available at: https://ssrn.com/abstract=4716839

that approximately in all of the topics, considering the overall sentiment index, the percentage of mildly and moderately negative sentiments is more than mildly and moderately positive ones. That is, most of the comments on ChatGPT tend to be negative but not extreme opinions.

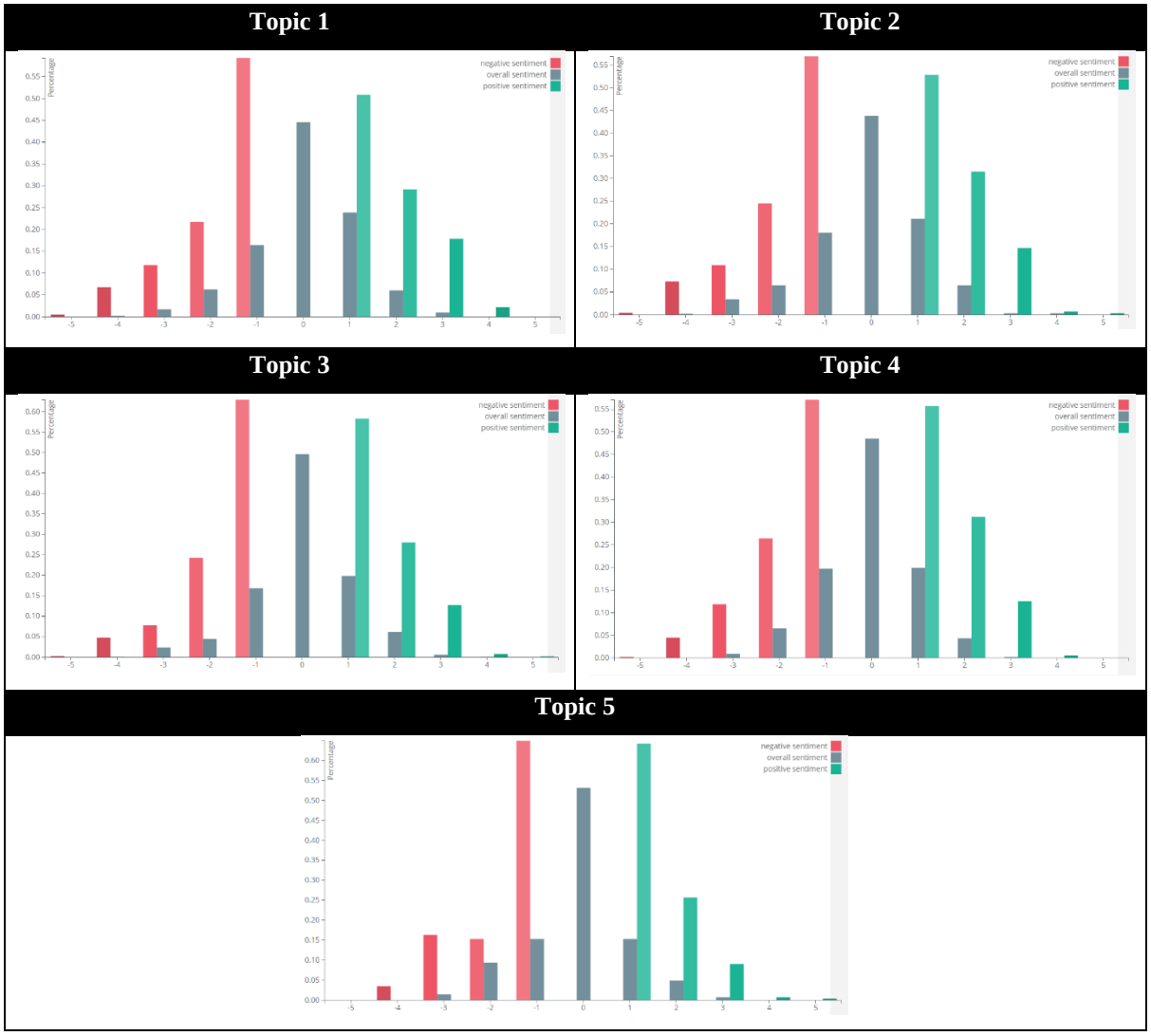


Figure 5. Negative, overall, and positive sentiments of topics

Table 1 presents the results of our VADER-based sentiment analysis of the two datasets. To understand the results, we focus on the compound sentiment score, which was explained in the method section, since it is the best measure if we seek a single unidimensional metric of sentiment for a specific body of text (Hutto, 2021). Accordingly, it is known that the posts and comments about ChatGPT on the Reddit social media platform are generally neutral while they tend to be a bit positive. Also, it can be seen that the sentiment of the second dataset is slightly

Electronic copy available at: https://ssrn.com/abstract=4716839

more positive than the first dataset. In addition, **Figure 6** depicts the distribution of compound scores for the first and second datasets that visually reflect the aforementioned understandings.

Table 1. Overall sentiment scores based on the VADER model

	Negative	Neutral	Positive	Compound
Dataset 1	0.096	0.702	0.202	**0.169**
Dataset 2	0.109	0.698	0.193	**0.171**

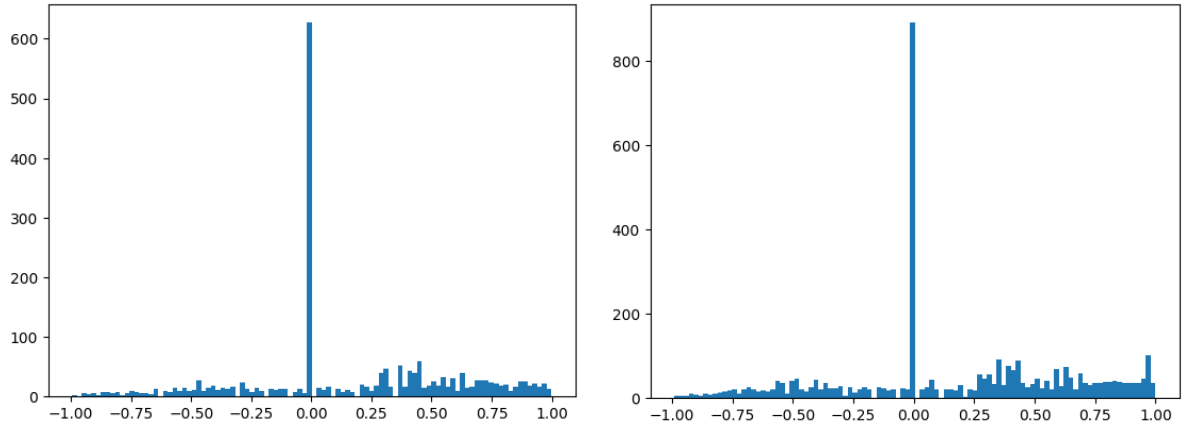


Figure 6. Distribution of compound scores for the first (left panel) and second (right panel) datasets

Discussion and conclusion

ChatGPT is one of the latest cutting-edge AI-based technologies that is disrupting many aspects of our world and setting a new bar for peoples' life and work. Accordingly, the present study attempted at understanding what is going on in the public's minds about this novel chatbot. To do so, we gathered all of the ChatGPT-related posts and comments that were left on the Reddit social media platform at two different time points in recent months. Next, we applied a topic modeling algorithm and two sentiment analysis techniques to analyze the data. As a result, the main public opinions and sentiments that are currently being discussed and expressed about ChatGPT were revealed. The topic modeling results indicated that the public's conversations about this technology surround around five main topics namely, *Human-ChatGPT Interaction*, *Credibility of ChatGPT Knowledge*, *Implications of ChatGPT for Life and Work*, *ChatGPT Technical Foundation*, and *Funny Chats with ChatGPT*. Besides, it could be inferred from the SentiStrength-based sentiment analysis results that people have both positive and negative opinions about this technology while most of them are mildly or moderately negative. As

Electronic copy available at: https://ssrn.com/abstract=4716839

ChatGPT is still in its infancy stage, it is interesting for everyone to interact with this tool for the sake of evaluating and knowing the real abilities of this AI-based chatbot. People are curious about how this tool works, what is the extent of its knowledge across different disciplines, and how this technology can digitally transform or facilitate their daily life and work. In this regard, although everyone appreciates the amazing capabilities of ChatGPT, they discuss that its knowledge is not flawless and it cannot substitute every aspect of our personal and professional lives. Based on what was obtained from our investigation, ChatGPT's users should be cautious about using this technology in their critical tasks due to its current limitations such as misunderstanding the input text, causing ethical and cybersecurity concerns, and generating plausible-sounding but inaccurate or nonsensical answers. In addition, the technology developers should try to redesign the tool to resolve its current shortcomings and offer the next versions with higher performance and accuracy. They also should try to publicly clarify the main mission, knowledge base, and technical foundations of ChatGPT and the next variants of such technologies. Doing so can increase the acceptance rate and perceived usability of the product and overcome the already-discussed concerns on the tool such as the politics and ideology behind ChatGPT (Deepak, 2023). The results of VADER-based sentiment analyses are also significantly important and bring thoughtful insights for the readers of the present study. The results indicate that although the Reddit users' perceptions were slightly negative toward ChatGPT a short time after its introduction, the opinions are gradually becoming positive at these days. It might be attributed to this fact that ChatGPT users have become more familiar with the real capabilities of this technology and understood how to usefully benefit from this tool instead of seeing the disruptions. It is consistent with this fact that the perception of emerging technologies experiences some phases such as emergence, growth, and maturity, so that potential users of the technology have usually unrealistic expectations of superior performance in the early phases (Pano and Kashef, 2020). The results of VADER sentiment analysis have also another important result. They give this opportunity to the researchers in the field of sentiment analysis to understand that the results of applying two different algorithms, the SentiStrength and VADER in our paper, on a dataset can be completely different and even in contrast with each other. Hence, we should try to mindfully use sentiment analysis algorithms given their specific applications and features. The pertinent research field is at its early stages, so there is a lot of room for further contributions. In this regard, in our opinion, *Human-ChatGPT Interaction* and *Implications of ChatGPT for Life and Work* are two very interesting areas for investigation by ChatGPT researchers. In these areas, they can explore some issues such as different drivers and barriers of adoption and acceptance along with

Electronic copy available at: https://ssrn.com/abstract=4716839

antecedents, nature, and consequences of various forms of usage. Also, scholars can examine the impacts of generative and conversational AI-enabled technologies for today and future variants of work and business models. Needless to say, this study has its own limitations. Firstly, in our analyses, we only used the data from one social media platform, so readers should be cautious about generalizing the findings. In this regard, it should be noted that the Reddit platform is usually utilized by relatively younger people (Li et al., 2023; Proferes et al., 2021), so our findings cannot be generalized to the whole population. Secondly, choosing an appropriate title for the topics and writing the relevant explanations have been subjective and limited to our current knowledge which is rooted in the present boundaries of the ChatGPT scholarship. Thirdly, a social media dataset can be subject to some biases or complexities such as ad-like and bot-generated content or sarcasm that are difficult to be detected and often lead to false positives or false negatives. To address these limitations, it is suggested to employ multiple social media platforms for data collection and analyze the data based on more advanced natural language processing techniques. Moreover, it is recommended to analyze other attributes of the social media data to provide other sorts of insights.

References

Adetayo AJ (2023) Artificial intelligence chatbots in academic libraries: the rise of ChatGPT. *Library Hi Tech News* 40(3): 18-21.

Ågerfalk PJ, Conboy K, Crowston K, et al. (2022) Artificial intelligence in information systems: state of the art and research roadmap. *Communications of the Association for Information Systems* 50(1): 420-438.

Baruffati A (2023) ChatGPT statistics 2023: trends and the future perspectives. *Gitnuxblog*. Retrieved from https://www.blog.gitnux.com/chat-gpt-statistics on 26 April 2023.

Bian J, Yoshigoe K, Hicks A, et al. (2016) Mining Twitter to assess the public perception of the "Internet of Things". *PloS One* 11(7): e0158450.

Bin-Hady WRA, Al-Kadi A, Hazaea A, et al. (2023) Exploring the dimensions of ChatGPT in English language learning: a global perspective. *Library Hi Tech* (ahead-of-print). DOI: 10.1108/LHT-05-2023-0200.

Electronic copy available at: https://ssrn.com/abstract=4716839

Birjali M, Kasri M, Beni-Hssane A (2021) A comprehensive survey on sentiment analysis: approaches, challenges and trends. *Knowledge-Based Systems* 226: 107134.

Blei DM (2012) Surveying a suite of algorithms that offer a solution to managing large document archives. *Communication of the ACM* 55(4): 77-84.

Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. *Journal of Machine Learning Research* 3(Jan): 993-1022.

Borg A, Boldt M (2020) Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications* 162: 113746.

Brandl R, Ellis C (2023) ChatGPT statistics 2023 all the latest statistics about OpenAI's chatbot. *Tooltester*. Retrieved from https://www.tooltester.com/en/blog/chatgpt-statistics on 19 August 2023.

Brans H, Scholtens B (2020) Under his thumb the effect of president Donald Trump's Twitter messages on the US stock market. *PloS One* 15(3): e0229931.

Burger B, Kanbach DK, Kraus S, et al. (2023) On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management* 26(7): 233-241.

Carvalho I, Ivanov S (2023) ChatGPT for tourism: applications, benefits and risks. *Tourism Review* (ahead-of-print). DOI: 10.1108/TR-02-2023-0088.

Cascella M, Montomoli J, Bellini V, et al. (2023) Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. *Journal of Medical Systems* 47(1): 1-5.

Chiru CG, Rebedea T, Ciotec S (2014) Comparison between LSA-LDA-lexical chains. In: *Proceedings of the international conference on web information systems and technologies*, Barcelona, Spain, 3-5 April 2014, pp. 255-262. Springer.

Choudhury A, Shamszare H (2023) Investigating the impact of user trust on the adoption and use of ChatGPT: survey analysis. *Journal of Medical Internet Research* 25: e47184.

Chow JC, Sanders L, Li K (2023) Impact of ChatGPT on medical chatbots as a disruptive technology. *Frontiers in Artificial Intelligence* 6: 1166014.

Chu MN (2023) Assessing the benefits of ChatGPT for business: an empirical study on organizational performance. *IEEE Access* 11: 76427-76436.

Electronic copy available at: https://ssrn.com/abstract=4716839

Cooper G (2023) Examining science education in ChatGPT: an exploratory study of generative artificial intelligence. *Journal of Science Education and Technology* 32: 444-452.

Cotton DR, Cotton PA, Shipway JR (2023) Chatting and cheating: ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International* (ahead-of-print). DOI: 10.1080/14703297.2023.2190148.

Deepak P (2023) ChatGPT is not OK! That's not (just) because it lies. *AI & Society* (ahead-of-print). DOI: 10.1007/s00146-023-01660-x.

Dowling M, Lucey B (2023) ChatGPT for (finance) research: the Bananarama conjecture. *Finance Research Letters* 53: 103662.

Dwivedi YK, Kshetri N, Hughes L, et al. (2023a) "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management* 71: 102642.

Dwivedi YK, Pandey N, Currie W, et al. (2023b) Leveraging ChatGPT and other generative artificial intelligence (AI)-based applications in the hospitality and tourism industry: practices, challenges and research agenda. *International Journal of Contemporary Hospitality Management* (ahead-of-print). DOI: 10.1108/IJCHM-05-2023-0686.

Elbagir S, Yang J (2019) Twitter sentiment analysis using natural language toolkit and VADER sentiment. In: *Proceedings of the international multiconference of engineers and computer scientists*, Hong Kong, China, 13-15 March 2019, pp. 12-16. Newswood Limited.

Esmaeili Givi M, Saberi MK, Talafidaryani M, et al. (2022) Assessment of the history and trends of "The Journal of Intellectual Capital": a bibliometrics, altmetrics and text mining analysis. *Journal of Intellectual Capital* 23(4): 864-912.

Farrokhnia M, Banihashem SK, Noroozi O, et al. (2023) A SWOT analysis of ChatGPT: implications for educational practice and research. *Innovations in Education and Teaching International* (ahead-of-print). DOI: 10.1080/14703297.2023.2195846.

Feldman R (2013) Techniques and applications for sentiment analysis. *Communications of the ACM* 56(4): 82-89.

Electronic copy available at: https://ssrn.com/abstract=4716839

Foroughi B, Senali MG, Iranmanesh M, et al. (2023) Determinants of intention to use ChatGPT for educational purposes: findings from PLS-SEM and fsQCA. *International Journal of Human-Computer Interaction* (ahead-of-print). DOI: 10.1080/10447318.2023.2226495.

Friedman A (2023) AI chatbot predicted to replace google in a couple of years. *Phonearena*. Retrieved from https://www.phonearena.com/news/ai-chatbot-could-replace-google-soon_id144120 on 26 April 2023.

Geerling W, Mateer GD, Wooten J, et al. (2023) ChatGPT has aced the test of understanding in college economics: now what?. *The American Economist* 68(2): 233-245.

Grant N, Metz C (2022) A new chat bot is a 'code red' for Google's search business. *New York Times*. Retrieved from https://www.nytimes.com/2022/12/21/technology/ai-chatgpt-google-search on 25 April 2023.

Grover P, Kar AK, Janssen M, et al. (2019) Perceived usefulness, ease of use and user acceptance of blockchain technology for digital transactions–insights from user-generated content on Twitter. *Enterprise Information Systems* 13(6): 771-800.

Gursoy D, Li Y, Song H (2023) ChatGPT and the hospitality and tourism industry: an overview of current trends and future research directions. *Journal of Hospitality Marketing & Management* 32(5): 579-592.

Haddara M, Hsieh J, Fagerstrøm A, et al. (2020) Exploring customer online reviews for new product development: the case of identifying reinforcers in the cosmetic industry. *Managerial and Decision Economics* 41(2): 250-273.

Hajiheydari N, Talafidaryani M, Khabiri S, et al. (2019) Business model analytics: technically review business model research domain. *Foresight* 21(6): 654-679.

Halaweh M (2023) ChatGPT in education: strategies for responsible implementation. *Contemporary Educational Technology* 15(2): ep421.

Haleem A, Javaid M, Singh RP (2022) An era of ChatGPT as a significant futuristic support tool: a study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations* 2(4): 100089.

Haque MU, Dharmadasa I, Sworna ZT, et al. (2022) "I think this is the most disruptive technology": exploring sentiments of ChatGPT early adopters using Twitter data. *arXiv*. DOI: 10.48550/arXiv.2212.05856.

Electronic copy available at: https://ssrn.com/abstract=4716839

Hassani H, Silva ES (2023) The role of ChatGPT in data science: how ai-assisted conversational interfaces are revolutionizing the field. *Big Data and Cognitive Computing* 7(2): 62.

Heidari A, Talafidaryani M (2021) Identifying the knowledge structure and research trend of the field of dynamic capabilities by topic modeling. *Journal of Strategic Management Studies* 12(45): 161-180.

Hutto C (2021) VADER sentiment analysis. *GitHub*. Retrieved from https://www.github.com/cjhutto/vaderSentiment on 12 August 2023.

Hutto C, Gilbert E (2014) Vader: a parsimonious rule-based model for sentiment analysis of social media text. In: *Proceedings of the international AAAI conference on weblogs and social media*, Ann Arbor, Michigan, USA, 1-4 June 2014, pp. 216-225. AAAI Press.

Ivanov S, Soliman M (2023) Game of algorithms: ChatGPT implications for the future of tourism education and research. *Journal of Tourism Futures* 9(2): 214-221.

Jelodar H, Wang Y, Yuan C, et al. (2019) Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications* 78: 15169-15211.

Jeon J, Lee S (2023) Large language models in education: a focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies* (ahead-of-print). DOI: 10.1007/s10639-023-11834-1.

Jeong B, Yoon J, Lee JM (2019) Social media mining for product planning: a product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management* 48: 280-290.

Kim J, Kim JH, Kim C, et al. (2023) Decisions with ChatGPT: reexamining choice overload in ChatGPT recommendations. *Journal of Retailing and Consumer Services* 75: 103494.

Kregel I, Koch J, Plattfaut R (2021) Beyond the hype: robotic process automation's public perception over time. *Journal of Organizational Computing and Electronic Commerce* 31(2): 130-150.

Lee H, Kang P (2018) Identifying core topics in technology and innovation management studies: a topic model approach. *The Journal of Technology Transfer* 43: 1291-1317.

Electronic copy available at: https://ssrn.com/abstract=4716839

Li S, Xie Z, Chiu DK, et al. (2023) Sentiment analysis and topic modeling regarding online classes on the reddit platform: educators versus learners. *Applied Sciences* 13(4): 2250.

Liu G, Ma C (2023) Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching* (ahead-of-print). DOI: 10.1080/17501229.2023.2240316.

Liu S, Wright AP, Patterson BL, et al. (2023) Using AI-generated suggestions from ChatGPT to optimize clinical decision support. *Journal of the American Medical Informatics Association* 30(7): 1237-1245.

Lund BD, Wang T (2023) Chatting about ChatGPT: how may AI and GPT impact academia and libraries?. *Library Hi Tech News* 40(3): 26-29.

Lund BD, Wang T, Mannuru NR, et al. (2023) ChatGPT and a new academic reality: artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology* 74(5): 570-581.

Medhat W, Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. *Ain Shams Engineering Journal* 5(4): 1093-1113.

Milmo D (2023) ChatGPT reaches 100 million users two months after launch. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app on 26 April 2023.

Montoyo A, Martínez-Barco P, Balahur A (2012) Subjectivity and sentiment analysis: an overview of the current state of the area and envisaged developments. *Decision Support Systems* 53(4): 675-679.

OpenAI (2023) ChatGPT: optimizing language models for dialogue. Retrieved from https://www.openai.com/blog/chatgpt on 11 April 2023.

Panda S, Kaur N (2023) Exploring the viability of ChatGPT as an alternative to traditional chatbot systems in library and information centers. *Library Hi Tech News* 40(3): 22-25.

Pano T, Kashef R (2020) A complete VADER-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19. *Big Data and Cognitive Computing* 4(4): 33.

Parsa A, Ebrahimzadeh MH (2023) ChatGPT in medicine; a disruptive innovation or just one step forward?. *The Archives of Bone and Joint Surgery* 11(4): 225-226.

Electronic copy available at: https://ssrn.com/abstract=4716839

Pavlik JV (2023) Collaborating with ChatGPT: considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator* 78(1): 84-93.

Prieto SA, Mengiste ET, García de Soto B (2023) Investigating the use of ChatGPT for the scheduling of construction projects. *Buildings* 13(4): 857.

Proferes N, Jones N, Gilbert S, et al. (2021) Studying reddit: a systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society* 7(2): 20563051211019004.

Rahman MM, Watanobe Y (2023) ChatGPT for education and research: opportunities, threats, and strategies. *Applied Sciences* 13(9): 5783.

Rai A, Constantinides P, Sarker S (2019) Next generation digital platforms: toward human-AI hybrids. *MIS Quarterly* 43(1): iii-ix.

Ratten V, Jones P (2023) Generative artificial intelligence (ChatGPT): implications for management educators. *The International Journal of Management Education* 21(3): 100857.

Ravi K, Ravi V (2015) A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-based Systems* 89: 14-46.

Rawas S (2023) ChatGPT: empowering lifelong learning in the digital age of higher education. *Education and Information Technologies* (ahead-of-print). DOI: 10.1007/s10639-023-12114-8.

Ray PP (2023) ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems* 3: 121-154.

Ruby D (2023) 30+ detailed ChatGPT statistics - users & facts. *Demandsage*. Retrieved from https://www.demandsage.com/chatgpt-statistics on 19 August 2023.

Rudolph J, Tan S, Tan S (2023) ChatGPT: bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching* 6(1): 1-22.

Ruksakulpiwat S, Kumar A, Ajibade A (2023) Using ChatGPT in medical research: current status and future directions. *Journal of Multidisciplinary Healthcare* 16: 1513-1520.

Electronic copy available at: https://ssrn.com/abstract=4716839

Saif H, Fernandez M, He Y, et al. (2014) Senticircles for contextual and conceptual semantic sentiment analysis of twitter. *Lecture Notes in Computer Science* 8465: 83-98.

Sallam M (2023) ChatGPT utility in health care education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare* 11(6): 887.

Short CE, Short JC (2023) The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights* 19: e00388.

Shoufan A (2023) Exploring students' perceptions of CHATGPT: thematic analysis and follow-up survey. *IEEE Access* 11: 38805-38818.

Strzelecki A (2023) To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments* (ahead-of-print). DOI: 10.1080/10494820.2023.2209881.

Sullivan M, Kelly A, McLaughlan P (2023) ChatGPT in higher education: considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching* 6(1): 1-10.

Taecharungroj V (2023) "What can ChatGPT do?" Analyzing early reactions to the innovative AI chatbot on Twitter. *Big Data and Cognitive Computing* 7(1): 35.

Talafidaryani M (2019) Exploring the application of text-mining in identifying the research directions of a study field: a study of dynamic capabilities field. Master's Thesis, Faculty of Management, University of Tehran, Tehran, Iran.

Talafidaryani M (2021) A text mining-based review of the literature on dynamic capabilities perspective in information systems research. *Management Research Review* 44(2): 236-267.

Talafidaryani M, Jalali SMJ, Moro S (2021) Digital transformation: toward new research themes and collaborations yet to be explored. *Business Information Review* 38(2): 79-88.

Talafidaryani M, Jalali SMJ, Moro S (2023) Tracing the evolution of digitalisation research in business and management fields: bibliometric analysis, topic modelling and deep learning trend forecasting. *Journal of Information Science* (ahead-of-print). DOI: 10.1177/01655515221148365.

Temsah MH, Aljamaan F, Malki KH, et al. (2023) ChatGPT and the future of digital health: a study on healthcare workers' perceptions and expectations. *Healthcare* 11(13): 1812.

Electronic copy available at: https://ssrn.com/abstract=4716839

Teubner T, Flath CM, Weinhardt C, et al. (2023) Welcome to the era of chatgpt et al. the prospects of large language models. *Business & Information Systems Engineering* 65(2): 95-101.

Thelwall M (2018) Can social news websites pay for content and curation? The SteemIt cryptocurrency model. *Journal of Information Science* 44(6): 736-751.

Thelwall M, Buckley K, Paltoglou G, et al. (2010) Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology* 61(12): 2544-2558.

Thorp HH (2023) ChatGPT is fun, but not an author. *Science* 379(6630): 313.

Tlili A, Shehata B, Adarkwah MA, et al. (2023) What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments* 10(1): 15.

Treen K, Williams H, O'Neill S, et al. (2022) Discussion of climate change on Reddit: polarized discourse or deliberative debate?. *Environmental Communication* 16(5): 680-698.

Uddin SJ, Albert A, Ovid A, et al. (2023) Leveraging ChatGPT to aid construction hazard recognition and support safety education and training. *Sustainability* 15(9): 7121.

Vaishya R, Misra A, Vaish A (2023) ChatGPT: is this version good for healthcare and research?. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* 17(4): 102744.

Wamba SF, Queiroz MM, Jabbour CJC, et al. (2023) Are both generative AI and ChatGPT game changers for 21st-Century operations and supply chain excellence?. *International Journal of Production Economics* 265: 109015.

Wei X, Croft WB (2006) LDA-based document models for ad-hoc retrieval. In: *Proceedings of the annual international ACM SIGIR conference on research and development in information retrieval*, Seattle, Washington, USA, 6-11 August 2006, pp. 178-185. ACM.

Wu L, Dodoo NA, Wen TJ, et al. (2022) Understanding Twitter conversations about artificial intelligence in advertising based on natural language processing. *International Journal of Advertising* 41(4): 685-702.

Electronic copy available at: https://ssrn.com/abstract=4716839

Yan D (2023) Impact of ChatGPT on learners in a L2 writing practicum: an exploratory investigation. *Education and Information Technologies* (ahead-of-print). DOI: 10.1007/s10639-023-11742-4.

Yang L, Wang J (2023) Factors influencing initial public acceptance of integrating the ChatGPT-type model with government services. *Kybernetes* (ahead-of-print). DOI: 10.1108/K-06-2023-1011.

Yau CK, Porter A, Newman N, et al. (2014) Clustering scientific documents with topic modeling. *Scientometrics* 100: 767-786.

Electronic copy available at: https://ssrn.com/abstract=4716839

Appendix

Selected contributions on ChatGPT

#	Author(s)	Year	Context	Type	Main contributions
1	Adetayo	2023	Library and information science	Non-empirical	The potentials of AI chatbots and ChatGPT along with their dangers in the context of academic libraries were investigated.
2	Bin-Hady et al.	2023	Education	Empirical	The dimensions in which ChatGPT helps students in their English language learning were explored through a qualitative study.
3	Burger et al.	2023	Management	Non-empirical	The recent trends in the use of AI-based tools like ChatGPT to assist with (management) research were explored, and a guide was proposed with a focus on systematic reviews.
4	Carvalho and Ivanov	2023	Tourism	Non-empirical	The uses, advantages, and risks of ChatGPT, in particular, and large language models, in general, for tourism were outlined.
5	Cascella et al.	2023	Health	Non-empirical	The potential applications and limits of ChatGPT as a large language model were highlighted in clinical and research scenarios.
6	Choudhury and Shamszare	2023	Health	Empirical	The impact of users' trust in ChatGPT on their intention and actual use of the technology was investigated through a survey.
7	Chu	2023	Business	Empirical	The predictors of user satisfaction and benefits related to the quality components of ChatGPT along with how the components influence organizational performance were examined through conducting a survey.
8	Cooper	2023	Education	Empirical	The possible implications and potentialities of ChatGPT for science education were explored by asking some questions from this chatbot and analyzing its answers.
9	Cotton et al.	2023	Education	Non-empirical	The opportunities and challenges, the potential risks and rewards, and the ethical issues with regard to the use of ChatGPT in higher education were discussed.
10	Dowling and Lucey	2023	Finance	Empirical	The degree to which ChatGPT can support the different stages of the financial research

Electronic copy available at: https://ssrn.com/abstract=4716839

#	Author(s)	Year	Context	Type	Main contributions
					process was explored based on experts' ratings of the platform's output.
11	Dwivedi et al.	2023b	Tourism	Non-empirical	The current challenges and practices along with a research agenda for implementing generative AI (such as ChatGPT) and similar tools in the tourism and hospitality sector were discussed.
12	Farrokhnia et al.	2023	Education	Non-empirical	ChatGPT's strengths and weaknesses along with its opportunities for and threats to education practice and research were argued.
13	Foroughi et al.	2023	Education	Empirical	Predictors of intention to use ChatGPT for educational purposes were investigated by conducting a survey among students.
14	Geerling et al.	2023	Economics	Empirical	ChatGPT's ability in passing standardized tests was evaluated by asking a series of questions from a standard economics test.
15	Gursoy et al.	2023	Tourism	Non-empirical	ChatGPT's benefits as well as the potential challenges and threats it can cause for the tourism and hospitality sector were comprehensively discussed.
16	Halaweh	2023	Education	Non-empirical	The potential concerns for using ChatGPT in education and strategies for the responsible use of this chatbot in this context were argued.
17	Hassani and Silva	2023	Data science	Non-empirical	The opportunities and challenges related to the application of ChatGPT in data science and how it can be employed to enhance different tasks in the field were discussed.
18	Ivanov and Soliman	2023	Tourism	Non-empirical	The implications of ChatGPT and similar large language model-based chatbots for tourism education and research were argued.
19	Jeon and Lee	2023	Education	Empirical	The relationship between ChatGPT and teachers was explored, and the complementary roles of each in education were identified through an experiment.
20	Kim et al.	2023	Consumer behavior	Empirical	It examined how people respond differently to recommendation options produced by ChatGPT by performing some experiments.
21	Liu and Ma	2023	Education	Empirical	The acceptance and usage of ChatGPT by English learners in their informal digital

Electronic copy available at: https://ssrn.com/abstract=4716839

#	Author(s)	Year	Context	Type	Main contributions
					learning activities were investigated by conducting a survey.
22	Liu et al.	2023	Health	Empirical	It was investigated through some experiments and surveys if ChatGPT can give useful recommendations for enhancing clinical decision support. Also, noninferiority was evaluated in comparison with human-generated recommendations.
23	Lund and Wang	2023	Library and information science	Non-empirical	The potential impacts, benefits, and ethical considerations of ChatGPT for the academia and libraries field were discussed.
24	Lund et al.	2023	Library and information science	Non-empirical	The potential impacts and ethical issues of ChatGPT in academia and scholarly research and publishing were discussed.
25	Panda and Kaur	2023	Library and information science	Non-empirical	The possibility of employing ChatGPT-based chatbot systems as an alternative to traditional systems in libraries and information centers was explored.
26	Pavlik	2023	Journalism and media	Non-empirical	Some reflections on the implications of generative AI and ChatGPT for journalism and media education were offered.
27	Prieto et al.	2023	Construction	Empirical	The applicability of ChatGPT to help in developing an automated construction schedule was explored by giving the chatbot the scope of a project and analyzing its responses in scheduling the project.
28	Rahman and Watanobe	2023	Education	Empirical	The opportunities and challenges of ChatGPT for education specifically, programming learning and teaching, were explored from the perspective of students and educators through some experiments and surveys.
29	Ratten and Jones	2023	Management	Non-empirical	It was discussed how ChatGPT as a form of generative AI can be employed by management educators.
30	Rawas	2023	Education	Non-empirical	The use of ChatGPT, its applications, pros and cons as well as implementation advice and future research directions in higher education were discussed.

Electronic copy available at: https://ssrn.com/abstract=4716839

#	Author(s)	Year	Context	Type	Main contributions
31	Rudolph et al.	2023	Education	Non-empirical	The implications of ChatGPT and its relevance for learning, teaching, and assessment in higher education were argued.
32	Ruksakulpiwat et al.	2023	Health	Non-empirical	The studies related to ChatGPT's utility in medical research, including its potential to give treatment, diagnosis, and medication were reviewed.
33	Sallam	2023	Health	Non-empirical	The usage of ChatGPT and its potential limitations in healthcare education, research, and practice were highlighted.
34	Short and Short	2023	Entrepreneurship	Empirical	It investigated how ChatGPT and similar tools can swiftly create and refine workable rhetoric useful for entrepreneurs by prompting it to generate entrepreneurial pitches mimicking a celebrity CEO archetype.
35	Shoufan	2023	Education	Empirical	Students' perceptions of ChatGPT and its pros and cons were explored through a thematic analysis and a follow-up survey.
36	Strzelecki	2023	Education	Empirical	The factors that can predict the adoption and use of ChatGPT among higher education students were examined through a survey.
37	Sullivan et al.	2023	Education	Empirical	The impacts of ChatGPT on higher education regarding the issues of academic integrity and student learning were explored by analyzing the news articles.
38	Taecharungroj	2023	Non-context-specific	Empirical	The general use cases of ChatGPT were investigated by analyzing the early tweets about this innovative AI chatbot.
39	Temsah et al.	2023	Health	Empirical	The perceptions of healthcare workers towards ChatGPT's utility along with the possible challenges for adopting and employing AI chatbots in healthcare were assessed by conducting a survey.
40	Tlili et al.	2023	Education	Empirical	The issues with using chatbots, specifically ChatGPT, in education were investigated by analyzing the tweets, interviews, and user experiences.
41	Uddin et al.	2023	Construction	Empirical	It was explored through an experiment if ChatGPT can help hazard recognition when

Electronic copy available at: https://ssrn.com/abstract=4716839

#	Author(s)	Year	Context	Type	Main contributions
					integrated into the curriculum of students pursuing a job in the construction sector.
42	Vaishya et al.	2023	Health	Empirical	The current position and the accuracy of the present version of ChatGPT in association with healthcare and medical research were argued based on some interactions with this AI chatbot.
43	Wamba et al.	2023	Operations and supply chain	Empirical	The benefits, challenges, and trends related to generative AI and ChatGPT in operations and supply chain management were surveyed among the practitioners.
44	Yan	2023	Education	Empirical	Students' behaviors and reflections in their interaction with ChatGPT in L2 writing classrooms were investigated through observations, surveys, and interviews.
45	Yang and Wang	2023	Administration	Empirical	The predictors of initial public acceptance of integrating the ChatGPT-type model with government services were explored by using an innovative method.

Electronic copy available at: https://ssrn.com/abstract=4716839